1. Correlation

Correlation

Claim Term	CMU's Construction	Marvell's Construction
correlation **'839 Patent Claims 11, 16, 19, 23** **'180 Patent Claim 6**	the degree to which two more items (here, noise in signal samples) show a tendency to vary together. **CMU Brf. at 19-20**	the expected (mean) value of the product of two random variables (*e.g.*, $E[r_i r_j]$, where r_i and r_j are signal samples at time I and time j, respectively). **Marvell Brf. at 17-21**

- The Dispute:
  - Should "correlation" be accorded its ordinary meaning in engineering and statistics (Marvell) or its lay meaning (CMU)?

Claim Language

- Correlation-sensitive branch metrics calculated from noise covariance matrices

US 6,201,839

15

outputting said branch weight.

11. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adapti… detecting a sequence of symbols stored on a h… magnetic recording device, comprising…

(a) performing a Virterbi-lik… plurality of signal s… lation sensitive branch…

(b) outputting a delayed deci…

(c) outputting a delayed signal…

(d) adaptively updating a plurality… matrices in response to said delayed… said delayed decisions;

(e) recalculating said plurality of correlat… branch metrics from said noise covariance… using subsequent signal samples; and

(f) repeating steps (a)–(e) for every new signal samp…

12. The method of claim 11 wherein said Virterbi-like sequence detection is performed using a PRML algorithm.

13. The method of claim 11 wherein said Virterbi-like sequence detection is performed using an FDTS/DF algorithm.

14. The method of claim 11 wherein said Virterbi-like sequence detection is performed using an RAM-RSE algorithm.

15. The method of claim 11 wherein said Virterbi-like sequence detection is performed using an MDFE algorithm.

16. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising the steps of:

(a) performing a Virterbi-like sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;

(b) outputting a delayed decision on the transmitted symbol;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and said delayed decisions;

(e) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and

(f) repeating steps (a)–(e) for every new signal sample.

17. The method of claim 16 wherein said channel has nonstationary noise.

18. The method of claim 16 wherein said channel has nonstationary signal dependent noise.

19. A detector circuit for detecting a plurality of data from a plurality of signal samples read from a recording medium, comprising:

a Viterbi-like detector circuit, said Viterbi-like detector circuit for producing a plurality of delayed decisions and a plurality of delayed signal samples from a plurality of signal samples;

a noise statistics tracker circuit responsive to said Viterbi-like detector circuit for updating a plurality of noise covariance matrices in response to said delayed decisions and said delayed signal samples; and

a correlation-sensitive metric computation update circuit responsive to said noise statistics tracker circuit for recalculating a plurality of correlation-sensitive branch metrics from said noise covariance matrices, said metrics from said noise covariance matrices, said branch metrics output to said Viterbi-like detector circuit.

… read signals, said detector circuit having a circuit for

11. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols stored on a high density magnetic recording device, comprising the steps of:

(a) performing a Virterbi-like sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;

(b) outputting a delayed decision on the recorded symbol;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and said delayed decisions;

(e) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and

(f) repeating steps (a)–(e) for every new signal sample.

See '839 Patent Claims 11, 16, 19, 23; '180 Patent Claim 6

Specification: Uses Mathematical Terminology

Euclidian branch metric. In the simplest case, the noise samples are realizations of independent identically distributed Gaussian random variables with zero mean and variance σ^2. This is a white Gaussian noise assumption. This implies that the correlation distance is L=0 and that the noise pdf s have the same form for all noise samples. The total ISI

'839 Patent 5:58-64

- random variables
- mean
- variance
- correlation distance
- covariance matrix
- expected values
- correlation-sensitive metric

The (L+1)x(L+1) matrix C_i is the covariance matrix of the data samples $r_i, r_{i+1}, \ldots, r_{i+L}$, when a sequence of symbols $a_{i-Kl}, \ldots a_{i+L+Kt}$ is written. The matrix c_i in the denominator of (11) is the L×L lower principal submatrix of $C_i=[c_i]$. The (L+1)-dimensional vector $\underline{N}_i$ is the vector of differences between the observed samples and their expected values when the sequence of symbols $a_{i-Kl}, \ldots, a_{i+L+Kt}$ is written, i.e.:

$$\underline{N}_i=[(r_i-m_i)(r_{i+1}-m_{i+1}) \ldots (r_{i+L}-m_{i+L})]^T \quad (12)$$

The vector $\underline{n}_i$ collects the last L elements of $\underline{N}_i$, $\underline{n}_i=[(r_{i+1}-m_{i+1}) \ldots (r_{i+L}-m_{i+L})]^T$.
With this notation, the general correlation-sensitive metric is:

$$M_i = \log det \frac{C_i}{det\ c_i} + \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i \quad (13)$$

'839 Patent 6:56-7:4

Specification: Equation 13

- The "correlation-sensitive" branch metric (Equation 13)

With this notation, the general correlation-sensitive metric is:

$$M_i = \log det \frac{C_i}{det\ \ c_i} + \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i \quad (13)$$

'839 Patent 6:66-7:4

- Other branch metrics not "correlation-sensitive"

Euclidian branch metric.

$$M_i = N_i^2 = (r_i - m_i)^2 \quad (8)$$

'839 Patent 5:59-6:14

Variance dependent branch metric.

$$M_i = \log\sigma_i^2 + \frac{N_i^2}{\sigma_i^2} = \log\sigma_i^2 + \frac{(r_i - m_i)^2}{\sigma_i^2} \quad (10)$$

'839 Patent 6:15-35

Specification: Branch Metric Computation

- Correlation is used in the computation of the branch metric (13)

- $E[\hat{C}(\hat{a})] = E[\underline{N}_i\underline{N}_i^T]$ calculates the expected value of the product of signal samples

US 6,201,839 B1

9

values . . . , 0.95, −1.05, 0, . . . Clearly, in this example, what should be a target value of 1 becomes either 1, 0.9, or 0.95 depending on the written sequence. Because mean values and not noisy samples are being considered, this deviation is due to nonlinearities in the system. There are two fixes for this problem. The first is to employ a nonlinear filter (neural network or Volterra series filter) that is capable of overcoming these nonlinear distortions. Although recently very

10

is expected to be achieved and $\beta(t)$ is made to approach a constant value. Initially, $\beta(t)$ is close to zero, to reflect the lack of a good prior estimate $\hat{C}(\hat{a})$, and to rely more on the data estimate. With time, $\beta(t)$ is increased and settles around a value close to 1.

The impact of the initial conditions in (22) decays exponentially fast. Hence, the algorithm (22) can be started with an arbitrary initial guess for the covariance matrix $\hat{C}(\hat{a})$, with

need for further mean corrections. The focus is shifted to tracking the noise covariance matrices needed in the computation of the branch metrics (13).

Assume that the sequence of samples $r_i, r_{i+1}, \ldots, r_{i+L}$ is observed. Based on these and all other neighboring samples, after an appropriate delay of the Viterbi trellis, a decision is made that the most likely estimate for the sequence of symbols $a_{i-K_l}, \ldots, a_{i+L+K_t}$ is $\hat{a}_{i-K_l}, \ldots, \hat{a}_{i+L+K_t}$. Here L is the noise correlation length and $K=K_l+K_t+1$ is the ISI length. Let the current estimate for the $(L+1)\times(L+1)$ covariance matrix corresponding to the sequence of symbols $\hat{a}_{i-K_l}, \ldots, \hat{a}_{i+L+K_t}$ be $\hat{C}(\hat{a}_{i-K_l}, \ldots, \hat{a}_{i+L+K_t})$. This symbol is abbreviated with the shorter notation, $\hat{C}(\hat{a})$. If the estimate is unbiased, the expected value of the estimate is:

$$E\hat{C}(\hat{a})=E[\underline{N}_i\underline{N}_i^T] \quad (21)$$

where $\underline{N}_i$ is the vector of differences between the observed samples and their expected values, as defined in (12).

'839 Patent 9:21-37

Prosecution History: Confirms Marvell's Construction

- Patent Office rejected CMU's claims over Huszar

> The Examiner rejected claims 11-22 as being anticipated by U.S. Patent No. 5,862,192 to Huszar et al. The Examiner stated that Huszar et al. "discloses a method for detecting a sequence that exploits the correlation between adjacent signal samples for

6/12/00 Amdt. at 8, '839 Patent File History (Marvell Exh. 22)

- CMU argued that correlation requires multiplying signal samples

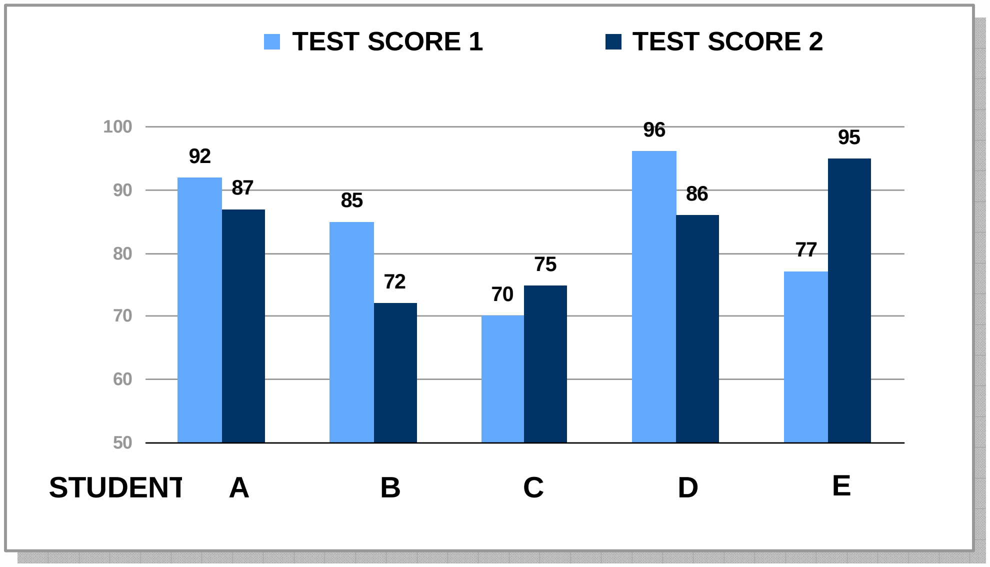
> Huszar et al. discloses branch metrics that are not correlation sensitive. Instead, the branch metrics of Huszar et al. are path metrics that have the form of (*See* Huszar et al., col. 8, equation 17):
>
> $$J = \Sigma_{\text{from } i=-\infty \text{ to } \infty} M_i$$
>
> where M_i is a branch metric of the form:
>
> $$M_i = [r_i(0) - y_i(0)]^2 + [r_i(1) - y_i(1)]^2$$
>
> Such a branch metric is not correlation sensitive, as claimed in independent claims 11, 16, and 19, which is evidenced by the fact that there is no term in the branch metric that corresponds to the correlation between $r_i(0)$ and $r_i(1)$, i.e. there is no term that involves multiplying $r_i(0)$ with $r_i(1)$. Thus, Huszar et al. does not disclose branch metrics that are correlation sensitive. Furthermore, Applicants submit that Huszar et al. does not disclose the use of noise covariance matrices. Because Huszar et al. does not disclose branch

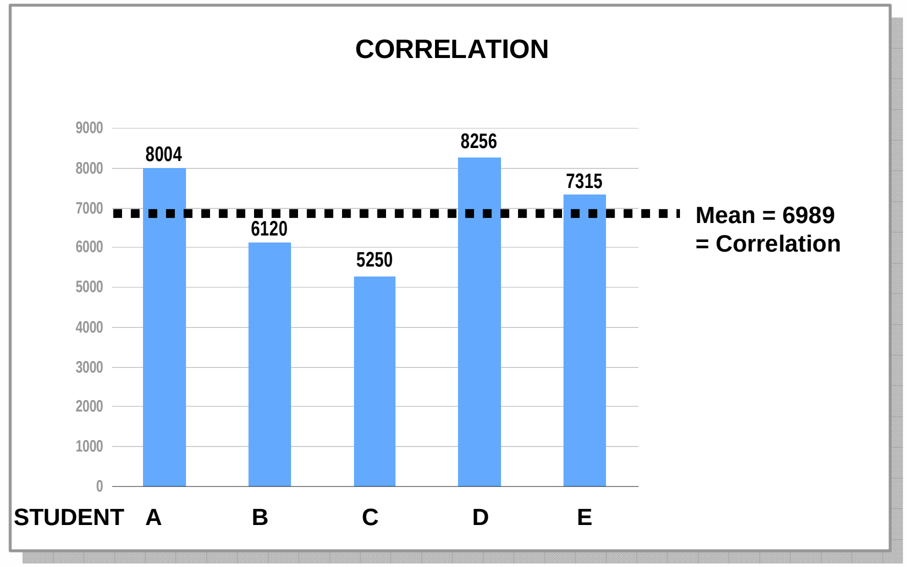
Id. at 8-9

Background: Statistics Definition

(Marvell Tutorial Slide 62)

CORRELATION the average of the pairwise products of test scores.





	TEST #1 SCORE	TEST #1 MINUS MEAN	DEVIATION SQUARED	TEST #2 SCORE	TEST #2 MINUS MEAN	TEST #1 x TEST #2 PRODUCT
STUDENT A	92	8	64	87	4	8004
STUDENT B	85	1	1	72	-11	6120
STUDENT C	70	-14	196	75	-8	5250
STUDENT D	96	12	144	86	3	8256
STUDENT E	77	-7	49	95	12	7315
	SUM 420		SUM 454	SUM 415		SUM 34945
	÷5 84		÷5 91	÷5 83		÷5 6989
	MEAN	DEVIATION	VARIANCE	MEAN	DEVIATION	CORRELATION

Extrinsic Evidence: Technical Treatises

- Marvell's construction is identical to statistical meaning:
  - ► *E*[*XY*] = the expected (mean) value of the product of two random variables X and Y

The second-order moment $m_{11} = E[XY]$ is called the *correlation* of X and Y. It is so important to later work that we give it the symbol R_{XY}.

Peebles, *Probability, Random Variables, and Random Signal Principles*, at 102 (1980) (Marvell Exh. 23)

X. In electrical engineering, it is customary to call the $j = 1$ $k = 1$ moment, $E[XY]$, the **correlation of X and Y**. If $E[XY] = 0$, then we say that **X and Y are orthogonal**.

Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, at 233 (1994) (Marvell Exh. 18)

  - ► *See also* Proakis Decl. at ¶¶ 30-31.

CMU's Reliance on General Dictionaries Fails

- CMU cites the Oxford English Dictionary

CMU Brf., at 21 n. 14

correlation (kɒrɪˈleɪʃən). [f. COR- + RELATION: cf. F. *corrélation*, and see CORRELATIVE.]

•
•
•

c. In *Statistics*, an interdependence of two or more variable quantities such that a change in the value of one is associated with a change in the value or the expectation of the others; also, the value of this as represented by a correlation coefficient. So *correlation coefficient* or *coefficient of correlation*: a number between −1 and 1 calculated so as to represent the linear interdependence of two variables or two sets of data; *spec.* the product-moment coefficient (see PRODUCT *sb.*[1]).

Compact Oxford English Dictionary (2d ed. 1987) (CMU Exh. 6)

- CMU truncated the definition that cited a "value"

CMU Truncates Dictionary Definitions

- CMU cites the Pocket Dictionary of Statistics

CMU Brf. at 22 n. 16

correlation– A general term denoting association or relationship between two or more variables. More generally, it is the extent or degree to which two or more quantities are associated or related. It is measured by an index called correlation coefficient. See also *intraclass correlation, Kendall's rank correlation, Spearman's rank correlation.*

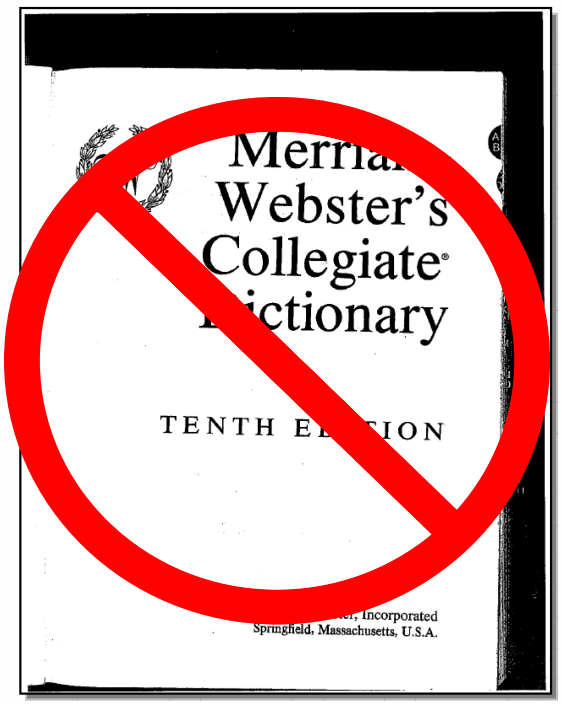
Pocket Dictionary of Statistics (2002) (CMU Exh. 10)

- Again omitting the quantitative reference

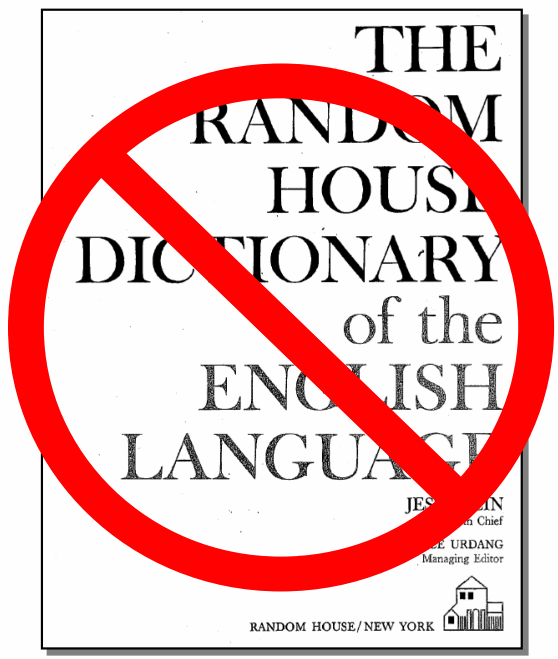
Improper Reliance on Dictionary Definitions

- "[I]t is inevitable that the multiple dictionary definitions for a term will extend beyond the construction of the patent that is confirmed by the avowed understanding of the patentee."
  Phillips, 315 F.3d at 1321-22, citing *Goodyear Dental Vulcanite Co. v. Davis*, 102 U.S. 222, 227, (1880).

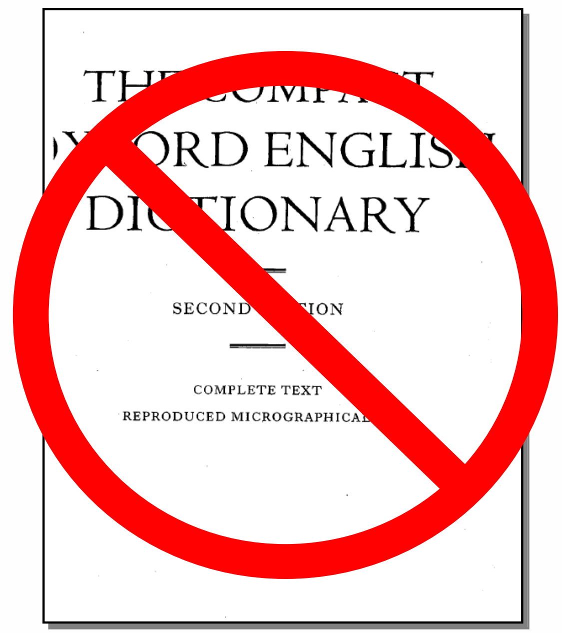
- "[A] general dictionary definition is secondary to the specific meaning of a technical term as it is used and understood in a particular technical field." *Hoechst Celanese Corp. v. BP Chems., Ltd.*, 78 F.3d 1575, 1580 (Fed. Cir. 1996).



CMU Exh. 3



CMU Exh. 4



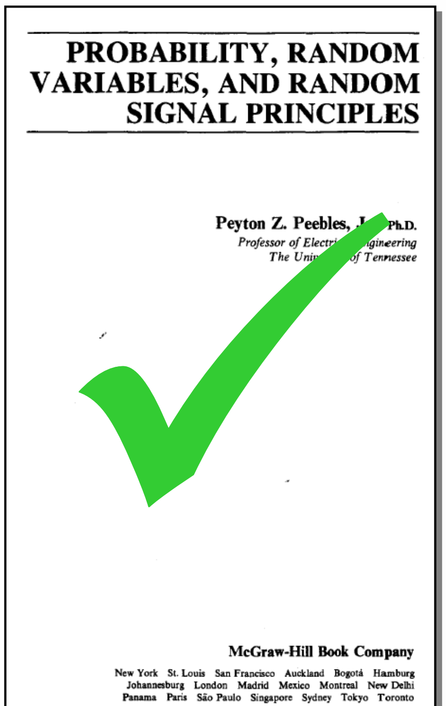
CMU Exh. 6

CMU Brf. at 21-22

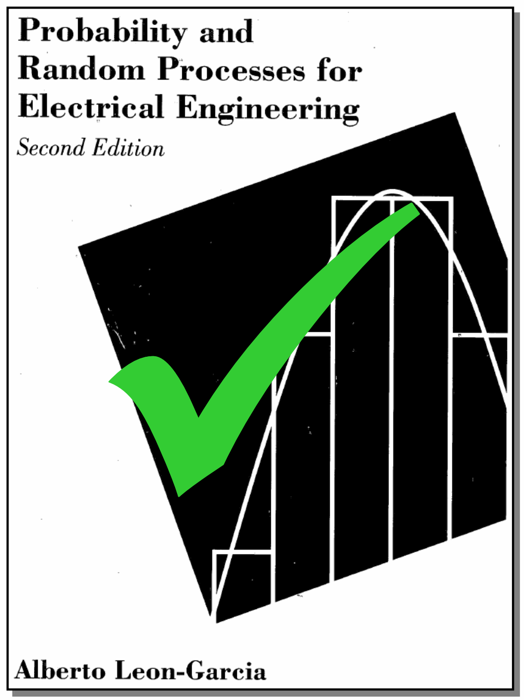
Technical Treatises Preferred Over Dictionaries

- "[T]echnical treatises [ ] constitute particularly strong sources of extrinsic evidence ... because they provide objective, contemporaneous, unbiased, and publicly available descriptions of how [terms are used] by those skilled in the art."

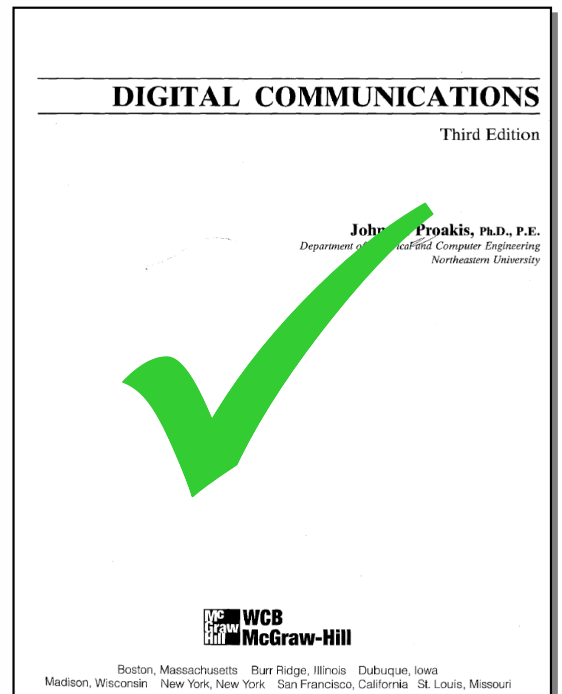
Osram GmbH v. Int'l Trade Comm'n, 505 F.3d at 1351, 1361 (Fed. Cir. 2007).



Marvell Exh. 24



Marvell Exh. 18



Marvell Exh. 36

Marvell Brf. at 18

CMU's Construction is (1) Vague and (2) Overbroad

- CMU's Construction: "*the degree to which two or more items show a tendency to vary together*"

1. Vague: "tendency to vary together" ?
   - May include statistics that are not covariance e.g. variance

variance– A **measure of variability** or **dispersion** of the values of a **data set** found by averaging the squared **deviations** about the **mean.** It is calculated by summing the squared

Sahai and Khurshid, *Pocket Dictionary of Statistics* (2002) (Marvell Exh. 16)

2. Overbroad: encompasses Euclidean branch metric that is not "correlation sensitive"

Euclidian branch metric.

$$M_i = N_i^2 = (r_i - m_i)^2 \qquad (8)$$

'839 Patent 5:59-6:14

CMU’s Incorrect Construction Covers Euclidean Metric

- “Euclidean branch metric” has noise samples that *vary together* [“identically” with “variance σ^2”]

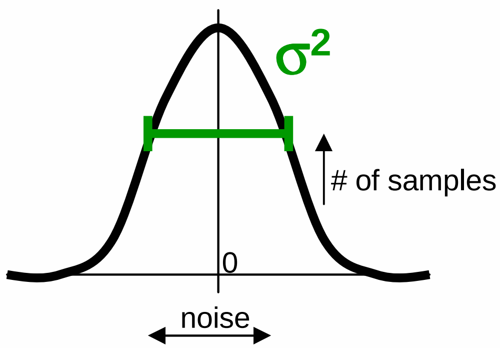
Euclidian branch metric. In the simplest case, the noise samples are realizations of independent identically distributed Gaussian random variables with zero mean and variance σ^2. This is a white Gaussian noise assumption. This

’839 Patent 5:59-62

$$M_i = N_i^2 = (r_i - m_i)^2 \quad (8)$$

’839 Patent 6:12-13

Gaussian noise



- CMU argued to the Patent Office:

$$M_i = [r_i(0) - y_i(0)]^2 + [r_i(1) - y_i(1)]^2$$

Such a branch metric is not correlation sensitive, as claimed in independent claims 11, 16,

6/12/00 Amdt. at 9, ’839 Patent File History
(Marvell Exh. 22)

CMU's "Disclosed Embodiment" Argument Fails

- CMU: "Marvell's proposed construction improperly excludes a disclosed embodiment"

CMU Brf. at 25

- Fails for two reasons:

1. No requirement that claims cover every embodiment
   - "The fact that a patent asserts that an invention achieves several objectives does not require that each of the claims be construed as limited to structures that are capable of achieving all of the objectives."

     Liebel-Flarshemi Co. v. Medrad, Inc., 358 F.3d 898, 908 (Fed. Cir. 2004).
   - "A Patentee may draft different claims to cover different embodiments."

     Intamin Ltd. v. Magnetar Technologies, Corp., 483 F.3d 1328, 1337 (Fed. Cir. 2007).
2. Marvell's construction covers the embodiments that correspond to the relevant claims

CMU's "Disclosed Embodiment" Argument Fails

- Marvell's construction covers the claims and embodiments that use "correlation"
  - ► **Group II Claims**
    - calculating "correlation-sensitive branch metrics" from "noise covariance matrices" (Eq.13; Fig. 3A)

> a correlation-sensitive metric computation update circuit responsive to said noise statistics tracker circuit for recalculating a plurality of correlation-sensitive branch metrics from said noise covariance matrices, said branch metrics output to said Viterbi-like detector circuit.

'839 Patent Claim 23; *see also* claims 11, 16, 19; '180 Patent Claim 6

- Other claims do not use "correlation"
  - ► **Group I Claims**
    - "method of determining branch metric values"
  - ► **Group III Claims**
    - "generating" a "branch weight" (weight w_i; Fig. 3B)

CMU's "Disclosed Embodiment" Argument Fails

- CMU concedes: "Figure 3A calculates a correlation … in one disclosed embodiment"

CMU Reply at 4

> FIG. 3A illustrates a block diagram of a branch metric computation circuit 48 that computes the metric M_i for a branch of a trellis, as in Equation (13). Each branch of the

'839 Patent 7:14-18

- ► Figure 3A computes Equation (13)
- ► Using Noise Covariance Matrix $\hat{C}$
- ► The $\hat{C}(\hat{a})$ estimate calculates correlation: $E[\hat{C}(\hat{a})] = E[N_i N_i^T]$
- ► **Marvell covers this embodiment**

> need for further mean corrections. The focus is shifted to tracking the noise covariance matrices needed in the computation of the branch metrics (13).
>
> Assume that the sequence of samples $r_i, r_{i+l}, \ldots, r_{i+L}$ is observed. Based on these and all other neighboring samples, after an appropriate delay of the Viterbi trellis, a decision is made that the most likely estimate for the sequence of symbols $a_{i-K_l}, \ldots, a_{i+L+K_t}$ is $\hat{a}_{i-K_l}, \ldots, \hat{a}_{i+L+K_t}$. Here L is the noise correlation length and $K=K_l+K_t+1$ is the ISI length. Let the current estimate for the $(L+1)\times(L+1)$ covariance matrix corresponding to the sequence of symbols $\hat{a}_{i-K_l}, \ldots, \hat{a}_{i+L+K_t}$ be $\hat{C}(\hat{a}_{i-K_l}, \ldots, \hat{a}_{i+L+K_t})$. This symbol is abbreviated with the shorter notation, $\hat{C}(\hat{a})$. If the estimate is unbiased, the expected value of the estimate is:
>
> $$E\hat{C}(\hat{a})=E[\underline{N}_i\underline{N}_i^T] \quad (21)$$
>
> where $\underline{N}_i$ is the vector of differences between the observed samples and their expected values, as defined in (12).

'839 Patent 9:21-37

CMU's "Disclosed Embodiment" Argument Fails

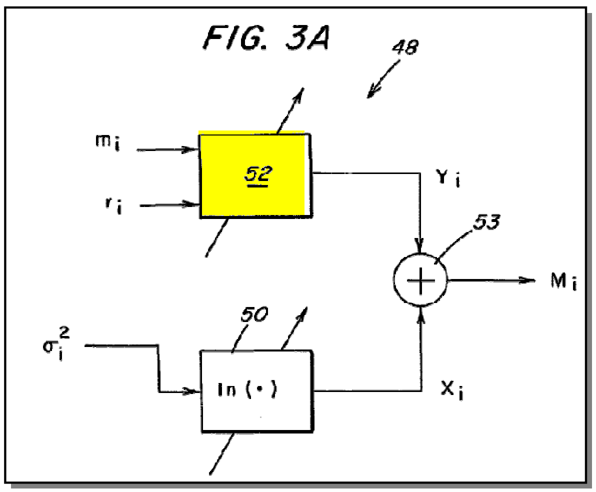
- McLaughlin: "an embodiment of which is shown in Figure 3B of the CMU Patents … without using or computing the expected value of the product of the signal samples."

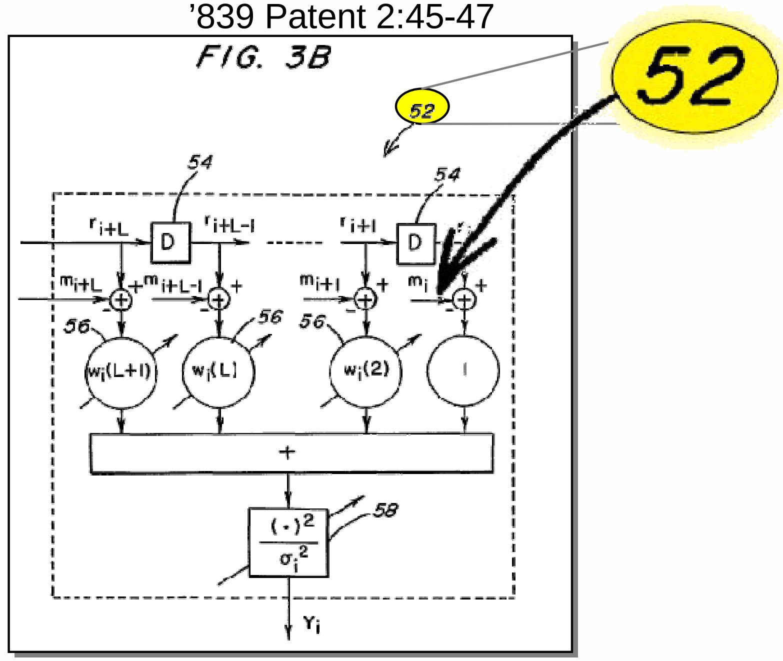
McLaughlin Decl. at ¶ 13

1. Fig. 3B only shows one implementation of Circuit 52 in Fig. 3A

> FIG. 3B is an illustration of an implementation of a portion of the branch metric computation module of FIG. 3A;

'839 Patent 2:45-47





2. Fig. 3B is covered by Group III claims (do not use "correlation")